

unenforceable in another. The *ProCD* case itself was based on a federal appeals court's reading of the Wisconsin Uniform Commercial Code.²¹⁷

A related jurisdictional issue raised by proponents is that even if state contract law is relatively consistent, many databases are marketed on a global scale. The contract laws of other countries tend to diverge more widely from the standard U.S. model, sometimes placing greater restrictions on freedom of contract based on each country's conceptions of public policy.²¹⁸

In addition, the enforceability of such contracts is not well settled, and has generated controversy. The *ProCD* case is the view of a single court of appeals, and may not prove to be the ultimate judicial word on the subject. Meanwhile, the ongoing project for the reform of the Uniform Commercial Code, administered by the American Law Institute (ALI) and the National Conference of Commissioners on Uniform State Laws (NCCUSL), includes a proposed Article 2B which would govern transactions in software and licenses in information. "Information" is defined to include data, databases, and "any intellectual property or other rights in information."²¹⁹ The issue of whether and to what extent such contracts can provide protection for data, or vary exceptions and limitations contained in the Copyright Act, is under debate.²²⁰

²¹⁷ *ProCD, Inc. v. Zeidenberg*, 908 F. Supp. 640, 651 (W.D. Wisc. 1996), *rev'd*, 86 F.3d (7th Cir. 1996). Although the license for a shrink-wrapped consumer item was held to be governed by U.C.C. art. 2 (sales), courts may not necessarily apply the U.C.C. to other licenses.

²¹⁸ See, e.g., *Turner Entertainment Co. v. Huston*, Court of Appeal of Versailles [France], Combined Civil Chambers, Decision No. 68, Roll No. 615/92 (Dec. 19, 1994) (as translated in 16 Ent. L. Rep. (March 1995)) (declining to enforce employment contract between U.S. film studio and U.S. director and screenwriter due to French public policy favoring moral rights). See also 1 MELVILLE B. NIMMER AND PAUL EDWARD GELLER, *INTERNATIONAL COPYRIGHT LAW AND PRACTICE* 198-99 (1993) (discussing enforcement of employment "work for hire" agreements outside United States generally).

²¹⁹ Article 2B-102(19) (May 5, 1997 Draft). See generally Raymond T. Nimmer, *Issues: Meeting the Information Age* (May 3, 1996) <<http://www.law.uh.edu/ucc/2b>>.

²²⁰ On May 19, 1997, ALI adopted an amendment to the current draft of section 2B-308, which deals with mass market licenses. The amendment reads: "In mass-market licenses, a term that is inconsistent with applicable provisions of the copyright law cannot become part of a contract" under the mass-market section. Transcript, ALI Annual Meeting (May 19, 1997), pp. 33-34. An earlier version of the amendment specifically prohibited terms that are inconsistent with section 102(b) of the Copyright Act, the codification

Finally, the argument has been made that contractual protection may not be optimal from the consumer's point of view. If relegated entirely to contractual self-help, database producers may make their products available only on license terms that are more restrictive than the terms that would be set by federal law. Such a trend may be developing already today, as some producers respond to their insecurity about legal protection after *Feist* by making databases available only with shrinkwrap licenses or on proprietary networks, and only upon terms barring many otherwise permissible uses.²²¹

Opponents question the seriousness of the privity problem. They assert that the chief value of many databases lies in their constant updating—especially those comprehensive databases that may not meet *Feist*'s creativity standard, and tend by their nature to be dynamic. If the producer of such a database suspects leakage, it can cut off access to the offending customer and block the current information flow that makes the database valuable.

In general, opponents stress that the law so far has confirmed the effectiveness of contractual means of protection, and that contracts today are relied on by many database producers. If the law should develop in a different direction, Congress could then consider the issue.

6. Misappropriation

Another existing form of protection for databases is provided by state common law under theories of misappropriation. The seminal case in this area dates back to 1918, when the Supreme Court held that the Associated Press (AP) had a claim against the International News Service

of the idea/expression dichotomy. If adopted into state law, this amendment might be read to overrule *ProCD*'s holding on this point and make it impossible for database producers to rely on contracts to limit the use of data in their databases.

On July 29, however, NCCUSL adopted a motion stating its belief that article 2B should not address the subject of this amendment, "but should adopt a position of neutrality on the issues which are being actively debated at federal and international levels," and suggested that ALI revisit its position.

²²¹ See *supra* section II.B.

(INS) to prevent it from copying news items from the war front gathered by AP at great trouble and expense, and scooping AP by making the items available to INS subscriber newspapers for advance publication.²²²

Although the *INS* decision was based on no-longer extant federal common law,²²³ it has been relied on over the years by various state courts in fashioning relief for similar conduct.²²⁴ *INS* was cited by the Supreme Court several times in the 1980s and '90s, including in *Feist*.²²⁵ Congress also referred to it in fashioning the preemption provision of the 1976 Copyright Act.²²⁶

²²² *International News Serv. v. Associated Press*, 248 U.S. 215 (1918).

²²³ See *Erie R.R. v. Tompkins*, 304 U.S. 64, 78 (1938).

²²⁴ See generally Douglas G. Baird, *Common Law Intellectual Property and the Legacy of International News Serv. v. Associated Press*, 50 U. CHI. L. REV. 411 (1983).

²²⁵ See *Feist*, 499 U.S. at 354 (stating that legal protection for facts "may in certain circumstances be available under theory of unfair competition"); *Carpenter v. United States*, 484 U.S. 19, 26 (1987); *San Francisco Arts & Athletics v. United States Olympic Comm.*, 483 U.S. 522, 532 (1987).

²²⁶ 17 U.S.C. § 301. See H.R. Rep. No. 1476, 94th Cong., 2d Sess. 132 (1976); S. Rep. No. 473, 94th Cong., 2d Sess. 116 (1976). The reference reads in full:

"Misappropriation" is not necessarily synonymous with copyright infringement, and thus a cause of action labeled as "misappropriation" is not preempted if it is in fact based neither on a right within the general scope of copyright as specified by section 106 nor on a right equivalent thereto. For example, state law should have the flexibility to afford a remedy (under traditional principles of equity) against a consistent pattern of unauthorized appropriation by a competitor of the facts (i.e., not the literary expression) constituting "hot" news, whether in the traditional mold of *International News Service v. Associated Press*, 248 U.S. 215 (1918), or in the newer form of data updates from scientific, business, or financial data bases. Likewise, a person having no trust or other relationship with the proprietor of a computerized data base should not be immunized from sanctions against electronically or cryptographically breaching the proprietor's security arrangements and accessing the proprietor's data. The unauthorized data access which should be remediable might also be achieved by the intentional interception of data transmissions by wire, microwave or laser transmissions, or by the common unintentional means of "crossed" telephone lines occasioned by errors in switching.

The proprietor of data displayed on the cathode ray tube of a computer terminal should be afforded protection against unauthorized printouts by third parties (with or without improper access), even if the data are not copyrightable . . .

The doctrine remained, however, somewhat ill-defined and uncertain in scope, as different courts applied it in different circumstances, sometimes without refined analysis.²²⁷

The misappropriation doctrine gained renewed clarity and authority earlier this year, when the Second Circuit decided *National Basketball Association v. Motorola, Inc.*²²⁸ In holding that a narrow form of common law misappropriation was not preempted by the Copyright Act, the Second Circuit delineated the elements of the surviving claim, and explained how it differed from copyright. According to the court, protection would be available under New York common law, without preemption, in the following circumstances:

(i) a plaintiff generates or gathers information at a cost; (ii) the information is time-sensitive; (iii) a defendant's use of the information constitutes free-riding on the plaintiff's efforts; (iv) the defendant is in direct competition with a product or service offered by the plaintiff; and (v) the ability of other parties to free-ride on the efforts of the plaintiff or others would so reduce the incentive to produce the product or service that its existence or quality would be substantially threatened.²²⁹

In the case before it, the court held that the National Basketball Association did not have a cause of action because it failed to show free-riding by the defendants or a sufficient competitive effect on the markets for its own products.²³⁰

Again, those seeking new federal protection acknowledge that the misappropriation doctrine is useful, but assert that it is insufficient for several reasons. First, they identify several potential shortcomings of the doctrine as elucidated by the Second Circuit. Whether or not the result in *Motorola* is appropriate, the court's analysis establishes the "hotness" or timeliness of the

²²⁷ See, e.g., *Metropolitan Opera Assoc., Inc. v. Wagner-Nichols Recorder Corp.*, 101 N.Y.S.2d 483 (Sup. Ct. 1950), *aff'd*, 279 App. Div. 63, 107 N.Y.S.2d 795 (1951).

²²⁸ 105 F.3d 841 (2d Cir. 1997).

²²⁹ *Id.* at 845.

²³⁰ *Id.* at 853-54.

data as a necessary element; apart from the question of how hot is “hot,” the value of many investment-rich databases may lie in the comprehensiveness of the collection of historical or timeless information. They also note that the commercial value of a database may be significantly harmed by unauthorized uses made by parties that are not in direct commercial competition, such as multiple uses by a member of the database’s intended audience or use by a commercial entity in preparing a related but distinct type of database. As under copyright law, they argue that the database producer should be able to protect its ability to exploit potential markets as well as those already being utilized.²³¹ Finally, the fifth element, relating to the reduction of incentives to produce, has been criticized as relating more to the degree of damage suffered by the database producer than to the nature of the wrongdoing.

More generally, proponents express concern that the tort of misappropriation is not well-defined or established in every state, and therefore leaves unclear where databases are protected and to what extent. The *Motorola* case is one decision in one Circuit, applying the law of one state; other state laws may be interpreted differently, and other courts may rule differently on the preemption issue. Proponents believe that both greater certainty and national uniformity are necessary for meaningful protection in today’s marketplace, especially in the on-line world. With inconsistent approaches in different states, difficult issues of choice of law and jurisdiction are likely to arise. These concerns are magnified in the international context.

Opponents of new protection view the misappropriation doctrine as a strong and effective means of protection, targeting with some precision the type of conduct most likely to cause meaningful commercial harm, while avoiding an impact on beneficial, public interest types of uses. As to the specific critiques of the doctrine described above, some argue that the “hotness” of the data is a reasonable criterion for protection, given the economic value of timely updates and the negatives to be weighed in the balance in protecting information. They also argue that as a policy

²³¹ Cf. 17 U.S.C. § 107(4) (in determining fair use, courts are to take into account effect of use on work’s potential market).

matter, limiting the legal claim to directly competitive markets is preferable, in order to avoid chilling the development of new, collateral database products. Competitors could then freely use information for different purposes, such as developing specialized niche databases. Or an entrepreneur could analyze historical financial data from the stock exchanges in order to predict future trends for investors.

As to the question of common law versus federal legislation, opponents assert that the courts are so far doing a reasonable job of interpreting and applying the doctrine of misappropriation. They urge that Congress should not step in prematurely where there does not yet appear to be a problem, but rather let the common law continue to develop. Some suggest that, if necessary, Congress instead consider amending section 301 of the Copyright Act to make clear that misappropriation claims are not preempted.

7. *Technological Protection*

In the increasingly important on-line environment, database publishers, like other creators of intangible materials, are looking to technological means to protect their products against unauthorized use. More and more sophisticated and effective forms are being developed today. Both owners and users cite such technological measures as critical elements of a workable system of protection, and at least a partial answer to the question of how to deal with the increased vulnerability to piracy in a digital world.

To proponents, such technological means of protection are necessary but not sufficient. The arguments made on both sides generally mirror the arguments that are made in the copyright context. On the one hand, technological protection has the potential to be extremely effective, easier and more economical to rely on than legal rights, and could obviate as a practical matter the need for additional legal protection; on the other hand, such protection is still in developmental stages, can be defeated by technological means of circumvention, does not prevent use of the database once someone has obtained an authorized copy in accessible form, and is effective primarily for databases in electronic form.

Some express concerns that technological protection could be too strong, making database producers completely invulnerable. They fear that producers, able to control every use made on-line, will impose stricter limits and permit less fair use.

8. *International Considerations*

One final aspect of the asserted need for new legislation is an outgrowth of the larger international context. Proponents point out that the market for databases, especially large and investment-intensive electronic databases, is global in scale. Accordingly, they are concerned not only about the level of protection in the United States, but in other countries as well.

In particular, they point to developments in Europe, where the recent directive on legal protection for databases effectively conditions protection for non-European Union databases on reciprocal protection in the given database's country of origin. In other words, an American database generally will not receive *sui generis* protection in EU member states unless U.S. law provides similar protection to databases.

Proponents argue that as of the beginning of 1998, when the directive's requirements take effect, American database producers will therefore be at a competitive disadvantage in Europe, one of the biggest markets for such works, as compared to their European counterparts. The latter will be able to control and profit from the use of their products, while the American producers will not. Related risks to market share are that U.S. producers will need to adopt more restrictive and less user-friendly contracts than their foreign competitors, and will not be able to operate safely in those jurisdictions where on-line or shrink-wrap contracts may not be respected. Concern is also expressed as to the effect on other countries outside the European Union of a failure to provide statutory protection in the United States, particularly those countries where piracy of U.S. works is a major problem today.

If, on the other hand, American database producers choose to avail themselves of the directive's alternate route to protection by establishing a commercial presence within the

European Union, proponents argue that the result will be a loss of jobs in the United States, with a corresponding detriment to the U.S. economy.

Opponents state that the United States should not follow Europe unless it is convinced that the European approach is a good idea. Rather, the United States should take the lead in establishing appropriate intellectual property policy, and seek to persuade the European Union and others to adopt our approach. This is particularly true, they argue, in areas relating to the use of government data, where the U.S. approach has historically differed from that of many European nations. There has long been controversy between the United States and Europe over appropriate treatment of such data, with the United States championing a policy of full and open access.²³²

Some opponents are concerned about potential negative international implications from enactment of new protection. They agree that markets are international in scope. Scientific research in particular increasingly involves international collaboration and the sharing of data collected globally. Several countries may participate in producing and maintaining a database, such as the database of DNA sequencing information created by the U.S., Japan and Europe. Research today requires the use of data sets from around the world. Science agencies caution that the United States should not send the wrong signal to other countries, and risk encouraging

²³² The U.S. government has been engaged for several years in espousing international agreements regarding full and open access to data. A multilateral policy to this effect was recently adopted in the World Meteorological Organization (WMO), with member countries agreeing to provide free and unrestricted exchange of meteorological and related data. WMO Policy and Practice for the Exchange of Meteorological and Related Data and Products Including Guidelines on Relationships in Commercial Meteorological Activities (WMO Resolution 40 (Cg-XII)) (1995). Numerous policy statements from international organizations and conferences, including the United Nations and the Organization for Economic Cooperation and Development, affirm this same goal in the context of other scientific disciplines as well.

governments to allow control of access to information, especially in a time of increasing budgetary constraints and corresponding commercialization of scientific data.²³³

Moreover, opponents express doubt that all member states of the European Union will have *sui generis* legislation in place by 1998, and believe it is uncertain what form such legislation might take. Even assuming the directive is fully implemented, they question whether it creates a real need for action. They assert that U.S. database producers will be no worse off in Europe than they are today, when *sui generis* protection does not exist, since they will merely fail to obtain the benefit of an *added* level of protection.²³⁴ They also point out that some U.S. producers already qualify for protection under the directive because they have a place of business in a member state, and those that do not can simply establish one.

Opponents question predictions of a meaningful competitive disadvantage. Rather, they believe any advantage to European producers will be only marginal. If a significant problem does arise, they argue, Congress can then respond.

Finally, opponents suggest that the directive's failure to provide national treatment may be challenged as an impermissible trade practice, inconsistent with existing treaty obligations, or as an inappropriate approach to intellectual property in a global marketplace.

C. Form of Any New Protection

If Congress determines that a need has been established for additional protection, the next question is what form that protection should take. Two basic models have been proposed, both in

²³³ Past proposals in the United States have made clear that there would be no protection for government data.. See H.R. 3531 § 3(c). Other countries, however, might choose to proceed differently (as they have in the area of copyright). See, e.g., Hong Kong Copyright Ordinance, Ord. No. 92 of 1997 §§ 182-186 (1997); United Kingdom Copyrights, Designs and Patents Act of 1988 §§ 163-167 (1988).

²³⁴ While the Directive will lower the level of copyright protection in some member states, it will raise the level in others. See discussion *supra* section IV.B. Accordingly, if *sui generis* protection is not provided to U.S. database producers in the European Union, they are likely to have less protection than today in some places and more in others.

the United States and in the course of debate over the directive in the European Union: (1) an exclusive property right; or (2) some form of unfair competition law, focusing on the nature of the conduct prohibited rather than providing ownership rights in particular subject matter.

The final version of the European directive adopts an exclusive property right model, as did the treaty proposals put on the table in WIPO last year and, at least arguably, the bill introduced in the 104th Congress. These approaches all provided database makers with certain specified rights in defined subject matter, lasting for a set period of time, transferable by contract, and subject to potential exceptions and limitations.

In contrast, an unfair competition model would not confer rights owned and enforceable against the world, but would make it unlawful to engage in conduct identified in some way as unfair.²³⁵ It would be closer to concepts contained in the Lanham Act, and embodied in the misappropriation doctrine set out in the *INS* and *Motorola* cases.

A federal misappropriation statute need not adopt every element of the state law claim outlined in *Motorola*, however, or in the same way. In the context of federal legislation, those elements that may be necessary to avoid preemption are not necessarily required, since in this context Congress itself would be determining where to draw the line between protection and free use. The issues would rather be the sufficiency of coverage of such legislation, and its compatibility with any constraints imposed by the Constitution.²³⁶

The choice between the two models has many potential ramifications. Depending on how it is drafted, an unfair competition model could obviate the need for definitions, for exceptions, or for a defined term of protection. The international consequences could also be quite different; in

²³⁵ Arguably, unfair competition principles were the true basis of the pre-*Feist* sweat of the brow directory cases. Commentators have noted the “reaping where one has not sown” language and rationale of many of the opinions. See, e.g., Jane C. Ginsburg, *Creation and Commercial Value: Copyright Protection of Works of Information*, 90 COLUM. L. REV. 1865, 1880-81 (1990). These cases generally involved commercial, competitive uses (although sometimes in related or potential, rather than directly overlapping, markets).

²³⁶ See *infra* section VII. G.

particular, an approach that differed significantly from the model of the European directive might not trigger reciprocal protection for U.S. databases in the member states. Finally, as discussed below, the two models may have differing constitutional implications.

Nevertheless, choosing one model or the other would only be the beginning. Much would turn on the precise delineation of either approach—how the scope of the rights are defined, or what conduct is proscribed. Many of the questions raised in the discussion of the remaining issues below would still need to be resolved.

As discussed above in section IV.B, the European Commission began with an unfair competition model, but ultimately adopted a property rights model in its Directive. The Commission has given several reasons for its change in approach, primarily: (1) the lack of established unfair competition laws in every country; (2) the need for producers to know what they own ahead of time, rather than waiting until someone engages in a use which a court finds wrongful; and (3) the commercial transferability of property rights.²³⁷

Proponents prefer the property rights model for these and other reasons. While their greatest concern may be unfair commercial conduct, and protection against free-riding, they point out that serious damage can be caused by an irresponsible user even without the elements of competition or profit. Proponents also are reluctant to rely on the existing state law misappropriation doctrine, given its checkered and ambiguous history. Their specific dissatisfactions with the *Motorola* formulation are described in more detail above.²³⁸

Among the opposing groups, and some neutral groups with specific concerns, there was a strong preference for the unfair competition model. While some felt that no need had been established for any legislation, and that it was preferable to let the courts continue to develop the common law, they were less uncomfortable with the former model than the latter. A number of

²³⁷ See Submission from the European Community and its Member States to the World Intellectual Property Organization on “An International Treaty on the Protection of Databases,” p. 2 (July 1997).

²³⁸ See *supra* section VII.B.5.

the concerns they expressed with regard to last year's proposals appeared to be ameliorated by such an approach. The more limited the formulation of unfair competition, and the closer to the Second Circuit's formulation in *Motorola*, the less objectionable some found it.

A few participants sought as much specificity as possible. They wished to avoid the uncertainty inherent in a general mandate to the courts to prevent conduct determined to be unfair. They urged that Congress take care not to adopt a law which would lead to litigation in every case over the legitimacy of the purpose for which data was taken. One scholar has suggested that users as well as producers would benefit from a clear statute establishing what types of use are and are not permissible, rather than continuing to rely on an ill-defined, potentially overbroad judge-made doctrine.²³⁹

D. Definitions

During the meetings, there was extensive discussion of the definitions used in the draft WIPO treaty and in last year's bill. In particular, participants focused on the definitions of "database," "substantial investment," and "substantial part" or its converse, "insubstantial part." While it was not assumed that the language from either of these proposals would be used this year, similar definitional issues may arise with any new proposals.

The definition of "database" raises the question of what exactly is the subject matter to be protected—a question that is integrally related to the nature and scope of the protection. All who commented on this question agreed that it is important to define the subject matter in such a way as not to sweep too broadly, and cover material that is not intended to be covered. Many pointed out that it is difficult to articulate a precise enough definition; some believed it to be impossible. There was substantial criticism of the definition of "database" in H.R. 3531 on this ground.

²³⁹ See Ginsburg, *supra* note 70. In this regard, it is interesting to note the reason for Justice Brandeis's dissent in *INS v. AP*. He believed that a remedy should exist for INS's conduct, but that such a claim should be provided by the legislature rather than by court-established common law. 248 U.S. at 267.

The following range of concerns was expressed: A broad definition, focusing on the collection of data in a systematic or accessible way, could be read to cover virtually everything in digital form. Even a motion picture or novel might qualify, as a systematically organized collection of 0s and 1s. Other collections of information, not ordinarily thought of as databases, might fall within the definition, such as on-line scientific discussions, scientific papers presenting research results, or an art historian's slide collection placed on-line. Computer programs, which are defined in the Copyright Act as a "set of statements or instructions," could be covered.²⁴⁰

Moreover, various building blocks of the Internet might be considered to qualify as databases, such as web sites, routing tables, domain name servers and interface specifications. If so, free access to these building blocks could be impaired, hindering interoperability and impeding the functioning of the Internet.

The video rental industry has a specific concern that the definition could cover videotapes, digital video disks, videogames or multimedia works generally, for example where a disk contains a movie combined with several previews or advertisements. Depending on how the form of protection was structured,²⁴¹ the result of including such items within the definition could be to establish for the first time in U.S. law a rental right for audiovisual works, making it impossible for companies like Blockbuster to continue their current rental business without obtaining licenses. It was suggested that one way to resolve this concern might be to require a minimum number of items to be collected in order for the collection to qualify as a database.²⁴²

²⁴⁰ See 17 U.S.C. § 101 (definition of "computer program"). In response to such concerns, the European Directive, the WIPO Draft Treaty and H.R. 3531 each included some form of a carve-out for computer programs. Database Directive art. 1(3); WIPO Draft Treaty art. 1(4); H.R. 3531 § 3(d).

²⁴¹ If, for example, rental was not included within the scope of any protection granted, this concern would not be a problem.

²⁴² *Cf.* Compendium, § 307.01 (establishing minimum numerical requirement of more than three items for work to be registered as compilation).

The scientific and educational communities in particular stressed the need to ensure that government data did not fall within the definition of protected subject matter. They believe this is even more important in the context of data than in the context of copyrightable subject matter. This goal could be accomplished through a specific exclusion, similar to that provided by H.R. 3531. The bill excluded databases produced by any government, in broader terms than the exclusion for U.S. government works in the Copyright Act,²⁴³ covering state and local governments as well as federal. Other possibilities would be an explicit exclusion of databases produced for the government by independent contractors as well as employees, or otherwise produced through the use of government funding, or databases produced by a private entity using data obtained from the government on an exclusive basis.²⁴⁴ In considering this issue, it should be borne in mind that some databases are created by international partnerships, and that treatment of government data may vary from country to country.

Proponents do not seek to protect government data itself, but stress the importance of providing incentives to private entities to create new, useful databases by investing in adding value to government data.²⁴⁵

The definition of “substantial investment” raises the issue of the criterion for protection. What kind of investment, and how much, should be required? The major concern expressed in the meetings related to the situation where someone takes a preexisting collection of data, and by adding limited value to it, obtains legal rights. This was identified as particularly problematic in the context of a private party adapting government or other public domain data in some way, involving no meaningful contribution of skill, judgment, or even effort, such as formatting or

²⁴³ 17 U.S.C. § 105.

²⁴⁴ See discussion of sole source databases *infra* section VII.F.

²⁴⁵ The OCLC database, for example, discussed *supra* in section III.B.1, adds value to Library of Congress catalogue records available through the Government Printing Office by providing codes identifying libraries around the world that have a given work in their collections.

adding page numbers, and then asserting control over its use. A number of participants stressed that significant added value should be required in order to obtain rights (and that the underlying information must remain available to others).

This issue is related to the question of duration, discussed below. If every new substantial investment qualifies a database for a new term of protection, the question of what constitutes a substantial investment is critical to how long protection will last. A low standard that requires only automated updating or reformatting could allow perpetual protection with little public benefit to justify it. On the other hand, a standard that is extremely high could obviate incentives for making expensive investments in researching and checking the information on a timely basis, and result in less useful databases.

The definition of “substantial part,” or its converse, “insubstantial part,” raises the issue of the scope of protection—i.e., what can be taken without implicating the legal rights. The European directive as well as the WIPO and legislative proposals last year provided protection against the taking of all or a substantial part of a database, excluding insubstantial portions from protection in themselves.²⁴⁶ This aspect of database protection is critical in ensuring that ordinary consumer or research use will be permissible without the need to obtain consent. Under all three prior models, a student could locate and extract from a database particular items of interest to him or her without implicating the producer’s rights.

In several of the Copyright Office meetings, concern was expressed that the terms used were vague, and that the taking of a single piece or small subset of data, if it were important or valuable, could be found by a court to be qualitatively substantial. The question was asked, for example, whether all sports scores from one particular game would be substantial. In addition, the WIPO draft treaty and H.R. 3531 each contained an exception to the general exclusion of insubstantial parts, in circumstances where those parts are accumulated in such a way as to affect

²⁴⁶ Directive art. 8(1); WIPO Draft Treaty art. 2(v); H.R. 3531 §§ 2 (defining “insubstantial part”) and 5(a) (1996).

the market for the database.²⁴⁷ Journalists and educators in particular were concerned about the possible impact of such an exception on news gathering and educational activities.

In response to the questions raised about the meaning of “substantial” and “insubstantial,” proponents point out that courts regularly interpret concepts of quantitative and qualitative “substantiality” and “materiality” in dealing with copyright and other bodies of law.

A few participants in the meetings suggested that some or all of the definitional questions could be avoided if an unfair competition model was chosen rather than a property rights model. By focusing on the nature of the conduct and the harm caused, rather than on the process of collecting the material itself, it might not be necessary to define precisely what material is and is not subject to protection.

E. Public Interest Uses

One fundamental concern was articulated by virtually all of the groups we met with that described themselves at least in part as database users. They identified certain activities with public interest elements that they urged should be allowed to continue without new restrictions on the ability to use data or new costs in doing so—primarily scientific, research and educational activities and news reporting. Each of these activities may span the range from non-profit to commercial in nature. Particular concern was expressed about the use of government and scientific data, sports statistics and financial data.

Analytically, there are various ways in which this concern could be addressed. One possibility relates to the form of protection chosen; depending on how it is articulated, a statute based on unfair competition is likely not to cover many such activities. If an exclusive property

²⁴⁷ See *supra* sections IV.C and V.

rights model is chosen instead, the scope of the rights granted could be drafted in such a way as to exclude such activities as appropriate.²⁴⁸

The exclusion from protection of insubstantial portions of a database helps but does not fully resolve the problem. While much education, research and reporting may rely on individual facts or small subsets of information, in some circumstances users need to extract substantial portions or all of a database in order to analyze its contents and draw conclusions. Thus, scientists often must analyze entire data sets in order to make findings and corroborate the research results of others, and may need to republish the background research for credibility. Public advocacy groups or reporters may need to examine substantial portions of a database to understand fully the scope of an issue.

Another possibility would be to provide an explicit exception or exceptions to cover those activities that Congress decides should be permitted without the need to obtain authorization. This could be accomplished through a broad, general exception similar to the fair use defense in copyright law; through detailed, specific exceptions more like the exceptions to a copyright owner's rights embodied in sections 108-121 of the Copyright Act; or through a combination of the two approaches.

A fair use-type approach provides several advantages. It is familiar and well-developed through judicial interpretation in the copyright context; it allows tremendous flexibility in adapting to particular factual circumstances;²⁴⁹ and it fits easily within the framework of guidelines for exceptions to rights within existing international intellectual property treaties.²⁵⁰ On the other

²⁴⁸ Cf. H.R. 3531 § 4(a) (right is violated only by use "that conflicts with the database owner's normal exploitation of the database or adversely affects the actual or potential market for the database").

²⁴⁹ The copyright fair use doctrine allows distinctions to be drawn between commercial and non-profit types of use, while recognizing that even the former may in appropriate circumstances qualify as fair. See, e.g., *Sega Enters. Ltd. v. Accolade, Inc.*, 977 F.2d 1510 (9th Cir. 1992).

²⁵⁰ See Berne Convention, art. 9(2); TRIPs, art. 13; WIPO Copyright Treaty, art. 10; WIPO Performances and Phonograms Treaty, art. 16. While the limitations on exceptions in these treaties may not apply to *sui generis* database protection, they represent a general approach toward exceptions that has

hand, there are disadvantages too. A fair-use type approach is unpredictable in its outcome in any given case, and therefore gives little certainty to users. It could also make the new form of protection appear more like copyright, raising the potential constitutional issue discussed below in section VII.G.

The specific exemptions approach presents the flip side of many of these advantages and disadvantages.

Finally, some have suggested the possibility of compulsory licenses for certain socially favored types of uses.²⁵¹ The rationale is that this would ensure the availability of data, while enabling the setting of a reasonable price. As a general rule, compulsory licenses are not favored in intellectual property law, which ordinarily relies on the marketplace, allowing rightholders freely to negotiate terms with users.²⁵² In some circumstances, however, Congress has found such licenses appropriate, typically where there is a new, struggling industry that Congress decides to assist, or some practical difficulty in achieving a negotiated solution.²⁵³

This leads to the question whether the marketplace can appropriately handle non-profit scientific and educational uses. Some databases are produced specifically for this market; others have both commercial and non-profit uses. As described in section II above, many database producers today engage in differential pricing. That is, they provide different terms for different types of uses, generally making databases available for much lower prices to nonprofit, scientific, library or educational users than to commercial users. In essence, the commercial users subsidize the non-commercial, allowing the producer to make a profit or at least cover costs. It is unclear

achieved international acceptance.

²⁵¹ See Reichman & Samuelson, *supra* note 145, at 146-148.

²⁵² See, e.g., Paul Goldstein, *Preempted State Doctrines, Involuntary Transfers and Compulsory Licenses: Testing the Limits of Copyright*, 24 U.C.L.A. L. REV. 1107, 1135-36 (1977) .

²⁵³ See, e.g., 17 U.S.C. §§ 111, 115 and 119. One approach taken has been to provide for compulsory arbitration if the parties cannot agree as to royalty rates and terms. See, e.g., 17 U.S.C. § 115(c)(3)(D).

whether or not enacting a new form of protection would alter this practice, or tend to raise prices overall, making access to data less affordable.

F. Duration

How long should protection last? All agree that, in theory, it should last just long enough to provide adequate incentives by allowing a fair return on investment. The difficulty lies in determining how long that period is. As with any form of intangible creation, it is complicated by the fact that different types of databases may need different terms to ensure a fair return. An extremely popular database of current and volatile data may recover costs in a much shorter time than an historical database requiring extensive research and appealing to a specialized audience. The challenge is to devise a term that works across the board, in order to encourage the production of all types of databases.

A number of possibilities have been suggested.²⁵⁴ The longest is the 25-year term proposed in last year's bill.²⁵⁵ The European directive requires a term of 15 years. The "catalogue rule" now in existence in some Nordic countries sets a term of 10 years. The "misappropriation" doctrine as set out in *INS* and *Motorola* suggests that protection may last as long as the data has value or as long as it is "hot"—i.e., new and timely. Such a term could vary for different databases, depending on the nature of the data, the particular market, and the state of communications technology. For example, stock prices today may be valuable or "hot" for only

²⁵⁴ Professors Reichman and Samuelson have suggested a combination of a short initial term, followed by a period where various compulsory licenses are in effect for different types of uses. See Reichman & Samuelson, *supra* note 145, at 147-48.

²⁵⁵ This term is considerably shorter than the term of protection for copyright. Most databases are works made for hire, and their copyright term would therefore last for seventy-five years from publication or one hundred years from creation, whichever expires first. 17 U.S.C. § 302(c).

fifteen minutes, while pre-television news from the front in World War I may have been “hot” for 24 hours or more.²⁵⁶

The discussion so far has dealt with the basic, initial term. The more difficult aspect of duration relates to changes made in a database, for example in the process of updating or verifying its contents. Proponents of legislation argue that a database should be protected as long as its producer continues to make substantial investments in maintaining it. For some databases, they report that producers spend many millions of dollars a year in updating and verifying the information they contain. They assert that there are equivalent public policy justifications for providing incentives to invest in keeping an existing database comprehensive, timely and accurate.

Last year’s bill dealt with this issue by providing that “any change of commercial significance” to an existing database, including by making additions, deletions or verifications, qualified the changed database for its own new term of protection.²⁵⁷ The WIPO draft treaty narrowed this language by adding the phrase “which constitute a new substantial investment” (the criterion for protection under the treaty).²⁵⁸ These provisions were controversial because they could be read to create a system of perpetual protection: as long as a database continued to be updated, new terms of protection could attach *ad infinitum*. This raised both policy and constitutional questions. Should Congress create a form of intangible property that could last forever? And would doing so violate the “limited times” restriction in the Copyright Clause of the Constitution?²⁵⁹

It also raised the question of what level of investment would be sufficient to qualify for an additional term. If the threshold is too low, there may be little justification for such extended

²⁵⁶ See *International News Serv. v. Associated Press*, 248 U.S. 215 (1918).

²⁵⁷ H.R. 3531 § 6(b).

²⁵⁸ Draft Database Treaty, arts. 1(1), 8(3). Cf. Database Directive, art. 10(3) (“Any substantial change . . . which would result in the database being considered to be a substantial new investment”).

²⁵⁹ U.S. CONST. art. I, § 8, cl. 8 (1789). See discussion *infra* section VII.G.1.

protection (particularly given the ease of making modifications in the digital age). If the requisite “substantial investment” is defined to be high enough, however, and the same level of investment is made that would qualify a new database for an initial term of protection, the argument has been made that protection should not be ruled out, simply because the comparable investment was made in updating and maintaining an existing database rather than creating a new one.

This treatment of changes made to existing databases parallels the treatment of changes to existing works of authorship in copyright law. When such a work is created today, it generally receives an initial term of protection measured by the life of the author plus fifty years.²⁶⁰ If someone lawfully makes changes to the work that in themselves qualify as creative authorship, the result is a derivative work, which is entitled to its own term of protection of life plus fifty.²⁶¹ The result is not perpetual protection, however; the Copyright Act states explicitly that this new term of protection is independent of and does not affect or enlarge the duration of any copyright in the preexisting work.²⁶²

One way to dispel the specter of perpetual protection might be to make explicit in any database legislation that the term of protection for the preexisting database is not extended when a new term attaches to a changed version. This would clarify the problem conceptually, and ensure that protection would expire in due course for the old version of the database. Thus, for example, if the West Publishing Company published a new version of its *Federal Reporter* series, with corrections to some older cases and incorporating new decisions, it would receive a full term of protection for the new version. Anyone would be free, however, to copy in its entirety the prior version of the series, once its set term of years had expired (if it was not protected by copyright).

²⁶⁰ 17 U.S.C. § 304.

²⁶¹ *Id.* §§ 101 (defining “derivative work”), 103(a).

²⁶² *Id.* § 103(b).

The remaining problem is a practical one. This solution will work for databases like the West reporters, to the extent that they are available in their original form. Databases available only on-line, however, may be constantly refreshed and not available to the public in their older form. Moreover, it may be impossible to determine which aspects of the database are new and which aspects were found in the prior version. The same problem exists today under copyright law. If the original work is not available, the fact that its term of protection has expired may not help a would-be user who has access only to a derivative work, particularly in situations where the preexisting material cannot easily be separated from the new matter.

It has been suggested that this issue too might be resolved by the choice of an unfair competition model rather than a property rights model. Again, the focus would be on fairness and commercial harm, rather than on the nature of the material taken. Protection could exist for as long as an investment of continued value was being taken unfairly.²⁶³

G. Sole Source Data

In theory, the answer to many of the concerns that have been expressed about restricting the availability of data is that, regardless of what model of protection is chosen, the database producer would not own the data in itself. The producer's rights would extend to its own particular database as an entity, but not the items collected in the database. In other words, anyone would remain free to obtain all of the same data from other sources. Thus, the legal protection would ensure that the database maker could protect the fruits of its investment in collecting and presenting data, but would leave others able to make their own collection of the

²⁶³ A comparison might be drawn to another branch of unfair competition, trademark law, under which rights exist as long as a mark continues to be used in commerce and to have value in identifying the source of the goods or services. See 15 U.S.C. §§ 1051, 1059 (Lanham Act §§ 1, 9).

same data. No participant at the meetings expressed disagreement with the concept of such a limitation, which could be explicitly stated in any legislation.²⁶⁴

Nevertheless, there are circumstances in which this answer alone may be unsatisfactory. When the data is not available elsewhere, the ability to prevent its extraction from the database may in effect amount to ownership of the data itself. The two prototypical examples of “sole source” data contained in a database are (1) government data provided to a private producer on an exclusive basis; and (2) data generated by the database maker itself. Included in the latter category are telephone subscriber information, sports statistics, and trading data from financial markets.²⁶⁵ Unless the producer chooses to make such data freely available, it is simply not possible for anyone else to obtain it independently.²⁶⁶

This is a complex issue, involving diverse types of databases and touching on a wide variety of policy implications. We present here some general points raised in the meetings as a preliminary stage in the analysis.

A variety of mechanisms have been proposed to deal with sole source databases. Broadly categorized, they are: exclusions from protection; compulsory licenses; and regulation through

²⁶⁴ *Cf.* H.R. 3531 § 5(b) (“[N]othing in this Act shall in any way restrict any person from independently collecting, assembling or compiling works, data or materials from sources other than a database subject to this Act”); defense of “independent creation” in copyright law. See, e.g., *Mazer v. Stein*, 347 U.S. 201, 218 (1954) (“Absent copying there can be no infringement of copyright”).

²⁶⁵ Other examples mentioned in the meetings included situations where the database producer may be the only entity in possession of the underlying information, for example where the original source no longer exists or has not retained the information; and situations where information may be available elsewhere but not in the “official” form demanded by users, such as sports league statistics or legal citations. *Cf.* H.R. 1584 and H.R. 1822, 104th Cong., 1st Sess. (1995) (barring, under certain circumstances, Federal and State courts and agencies from requiring a single citation form in which copyright subsists).

²⁶⁶ We do not suggest that all of the examples given should be treated in the same way. Different types of sole source data may raise different considerations, particularly with regard to the degree of justification for protection and the degree of need for access. Sports statistics in particular may be available as a practical matter through a variety of sources because the games are widely disseminated by television and radio broadcasts. See, e.g., *NBA v. Motorola, Inc.*, 105 F.3d (2d Cir. 1997) (scores obtained by defendant from television and radio).

other bodies of law such as antitrust or industry-specific government oversight. A combination of these approaches could also be considered, allowing greater fine-tuning to the nature of the database and its market.

A complete exclusion from protection is the most drastic approach, as it will result in a loss of the legal incentive to produce the database in question. This approach therefore implies a policy decision not to provide such an incentive for that type of database, and the absence of suitable, less drastic alternatives to ensure the availability of data.

The least controversial case for an exclusion from protection is the category of government data made available to the database producer on an exclusive basis. This issue implicates general U.S. policies about the conditions on which government data is made available to the public. Under current law, federal agencies are generally prohibited from entering into exclusive or restricted agreements for distribution of public information “that interferes with [its] timely and equitable availability to the public.”²⁶⁷ Nevertheless, the statute contains some exceptions, and other countries have different rules. The policy favoring free access to government data could be undermined if a single entity were permitted to control access through its database, with the public unable to obtain the data directly from the government or any third-party provider.²⁶⁸ This result could be avoided by broadening any statutory exclusion of databases

²⁶⁷ 44 U.S.C. § 3506(d)(4).

²⁶⁸ Under current law the data usually remains available from the government, but without the added value provided by the private sector producer. Government contracts for the publication of information generally require a continued non-exclusive license for the government to use the information and make it available to others, and may also require the producer to provide the information to the government in a more accessible form (e.g., automated). For example, the catalogue entries for copyright registrations from 1978 to date are available on-line through the Library of Congress. Those records are also available in a more accessible, user-friendly form from DIALOG Information Services, Inc., which provides a powerful search engine to its users.

created by a government entity to encompass databases created from government data that has been made available on exclusive terms to the database producer.²⁶⁹

The compulsory license approach may be seen as a middle ground, allowing producers to benefit financially from the use of their products but removing their ability to control the nature or price of the use. As discussed above, however, compulsory licenses are generally disfavored in intellectual property law, and adopted only as a last resort in circumstances where the free market does not function well. The idea of a compulsory license for sole source databases was proposed in Europe in the initial stages of the database directive, but abandoned as part of an overall compromise when it proved controversial.²⁷⁰

The third possibility is to deal with this issue as a question of appropriate government control of business activities. This could be done through the application of antitrust law generally, or through regulation of a particular industry, such as through the Federal Communications Commission for the telecommunications industry or through the Securities Exchange Commission for securities markets. These are areas where Congress has determined that a regulatory scheme is advisable in order to balance the interests of the industries and the public.

An example of the antitrust approach is the *Magill* decision in the European Court of Justice, which held that television broadcasters could not rely on their compilation copyrights to prevent the copying of self-generated programming information by others wishing to publish

²⁶⁹ The related issue of how to treat arrangements that are exclusive not as a legal matter but *de facto* is discussed below.

²⁷⁰ See discussion *supra* section IV.B.